

In [49]:

```
import numpy as np
x2 = np.array( [ [12, 5, 2, 4] , [7, 6, 8, 8] , [1, 6, 7, 7] ] )
print(x2[0, :])
```

```
[12  5  2  4]
```

In [50]:

```
import numpy as np
d = np.zeros(4, dtype={'names':('name', 'age', 'weight'),'formats':('U10', 'i4', 'f8')})
print(d.dtype)
```

```
[('name', '<U10'), ('age', '<i4'), ('weight', '<f8')]
```

#PART B QUESTION 5

How to create an array using Numpy packages ? Write the suitable code for the following function.

a).One Dimensional Array b).Two dimensional Array Declaration c) Boolean function d) Shape

In [51]:

#PART B QUESTION 5

```
import numpy as np
np.random.seed(0) # seed for reproducibility
x1 = np.random.randint(10, size=6) # One-dimensional array
x2 = np.random.randint(10, size=(3, 4)) # Two-dimensional array
print(x1)
print(x2)
print("x2 ndim: ", x2.ndim)
print("x2 shape:", x2.shape)#shape of an array
print("x2 size: ", x2.size)
print(x2==8)#Boolean Function
```

```
[5 0 3 3 7 9]
[[3 5 2 4]
 [7 6 8 8]
 [1 6 7 7]]
x2 ndim:  2
x2 shape: (3, 4)
x2 size:  12
[[False False False False]
 [False False  True  True]
 [False False False False]]
```

#PART B QUESTION 6

Let us assume that “array([50, 71,92, 63, 24, 15, 26, 47, 68, 89, 10, 1, 2, 58, 66, 5, 76, 7, 38, 19])”. Write the code and output for the following function a). Shape b)convert to Two-Dimensional array c). Conversion to Float d).Display the values using slicing or indexing

In [56]:

#PART B QUESTION 6

```
import numpy as np
arr=np.array([ 50, 71,92, 63, 24, 15, 26, 47, 68, 89, 10, 1, 2, 58, 66, 5, 76, 7, 38, 19])
print(arr)
print("\n")
print("arr shape: ",arr.shape)#shape of an array
print("\n")
arr1 = arr.reshape(4,5)#1 dimensional array to 2 dimensional array
print(arr1)
print("\n")
a = arr1.astype(np.float64)#convert to float
print(a)
print("\n")
print(a[2,3])#display values using indexing
print("\n")
print(a[1:,:])#display values using slicing
print("\n")
```

```
[50 71 92 63 24 15 26 47 68 89 10  1  2 58 66  5 76  7 38 19]
```

```
arr shape: (20,)
```

```
[[50 71 92 63 24]
 [15 26 47 68 89]
 [10  1  2 58 66]
 [ 5 76  7 38 19]]
```

```
[[50. 71. 92. 63. 24.]
 [15. 26. 47. 68. 89.]
 [10.  1.  2. 58. 66.]
 [ 5. 76.  7. 38. 19.]]
```

```
58.0
```

```
[[15. 26. 47. 68. 89.]
 [10.  1.  2. 58. 66.]
 [ 5. 76.  7. 38. 19.]]
```

#PART B QUESTION 8 Create a pandas series with three columns namely “Student Name, Dept and Mark” with sample values. Write the suitable code for the following pandas function.

a). Insert new values b).Delete the student data if mark<80 c). Display the last two rows d). Change any one student data e) Apply slicing or indexing

In [53]:



#PART B QUESTION 8

```
import pandas as pd
students = [['Amy', 'CSE', 98],
            ['joy', 'ECE', 75],
            ['roy', 'EEE', 70],
            ['Andy', 'CSE', 77],
            ['May', 'ECE', 95]]
df = pd.DataFrame(students, columns=['Name', 'Department', 'Mark'])
print(df)
print("\n")
df.loc[len(df.index)] = ['Mary', 'EEE', 93] #Insert new values
print(df)
print("\n")
print(df[df.Mark < 80]) #Delete the student data if mark < 80
print("\n")
print(df[4:]) #Display the last two rows
print("\n")
df = df.replace(['May'], 'Michael') #Change any one student data
print(df)
print("\n")
print(df.loc[2]) #Applying indexing
print("\n")
print(df[2:4]) #Applying Slicing
print("\n")
```

	Name	Department	Mark
0	Amy	CSE	98
1	joy	ECE	75
2	roy	EEE	70
3	Andy	CSE	77
4	May	ECE	95

	Name	Department	Mark
0	Amy	CSE	98
1	joy	ECE	75
2	roy	EEE	70
3	Andy	CSE	77
4	May	ECE	95
5	Mary	EEE	93

	Name	Department	Mark
1	joy	ECE	75
2	roy	EEE	70
3	Andy	CSE	77

	Name	Department	Mark
4	May	ECE	95
5	Mary	EEE	93

	Name	Department	Mark
0	Amy	CSE	98
1	joy	ECE	75
2	roy	EEE	70
3	Andy	CSE	77
4	Michael	ECE	95
5	Mary	EEE	93

```
Name      roy
Department  EEE
Mark      70
Name: 2, dtype: object
```

	Name	Department	Mark
2	roy	EEE	70
3	Andy	CSE	77

#PART B QUESTION 13

Let us assume that the dataset consists of five columns namely "Rollno, Name, Dept, Mark and Location". Write the code for the following function

A. How to upload the dataset?.

- B. Display the first 5 and last 5 rows.
- C. Display the information of Name column
- D. Change the column name Rollno as Regno
- E. Replace the null values using any one method
- F. Delete the column "Mark"
- G. Display the Mark>80 and <95
- H. Insert new values in 10th row.
- I. Insert new column as GPA.
- J. Find the sum of null values
- K. Create subset for first 3 columns

In [88]:



```
#PART B QUESTION 13
#to upload the dataset
import pandas as pd
data = pd.read_csv('STUDENT DATASET IAT.csv')
data.head(10)
```

Out[88]:

	ROLL NO	NAME	DEPARTMENT	MARK	LOCATION
0	1	MAY	CSE	98.0	CHENNAI
1	2	ANDY	EEE	54.0	TRICHY
2	3	ANDREW	ECE	71.0	MADURAI
3	4	MICHAEL	ECE	62.0	THENI
4	5	MARY	EEE	85.0	COIMBATORE
5	6	JOY	CSE	90.0	VELLORE
6	7	ROY	CSE	NaN	THANJAVUR
7	8	AMY	IT	45.0	KANYAKUMARI
8	9	ANNIE	EEE	80.0	CHENNAI
9	10	ESTHER	ECE	90.0	TRICHY

In [89]:

```
#Display the first 5 and last 5 rows.
print(data[:5])
print("\n")
print(data[5:])
```

	ROLL NO	NAME	DEPARTMENT	MARK	LOCATION
0	1	MAY	CSE	98.0	CHENNAI
1	2	ANDY	EEE	54.0	TRICHY
2	3	ANDREW	ECE	71.0	MADURAI
3	4	MICHAEL	ECE	62.0	THENI
4	5	MARY	EEE	85.0	COIMBATORE

	ROLL NO	NAME	DEPARTMENT	MARK	LOCATION
5	6	JOY	CSE	90.0	VELLORE
6	7	ROY	CSE	NaN	THANJAVUR
7	8	AMY	IT	45.0	KANYAKUMARI
8	9	ANNIE	EEE	80.0	CHENNAI
9	10	ESTHER	ECE	90.0	TRICHY

In [90]:

```
#Display the information of Name column
display(data.NAME)
```

```
0    MAY
1    ANDY
2    ANDREW
3    MICHAEL
4    MARY
5    JOY
6    ROY
7    AMY
8    ANNIE
9    ESTHER
Name: NAME, dtype: object
```

In [91]:

```
#Change the column name Rollno as Regno
data.rename(columns = {'ROLL NO':'Regno'}, inplace = True)
print(data)
```

	Regno	NAME	DEPARTMENT	MARK	LOCATION
0	1	MAY	CSE	98.0	CHENNAI
1	2	ANDY	EEE	54.0	TRICHY
2	3	ANDREW	ECE	71.0	MADURAI
3	4	MICHAEL	ECE	62.0	THENI
4	5	MARY	EEE	85.0	COIMBATORE
5	6	JOY	CSE	90.0	VELLORE
6	7	ROY	CSE	NaN	THANJAVUR
7	8	AMY	IT	45.0	KANYAKUMARI
8	9	ANNIE	EEE	80.0	CHENNAI
9	10	ESTHER	ECE	90.0	TRICHY

In [92]:



```
#Replace the null values using any one method  
data.fillna(100, inplace=True)  
data
```

Out[92]:

	Regno	NAME	DEPARTMENT	MARK	LOCATION
0	1	MAY	CSE	98.0	CHENNAI
1	2	ANDY	EEE	54.0	TRICHY
2	3	ANDREW	ECE	71.0	MADURAI
3	4	MICHAEL	ECE	62.0	THENI
4	5	MARY	EEE	85.0	COIMBATORE
5	6	JOY	CSE	90.0	VELLORE
6	7	ROY	CSE	100.0	THANJAVUR
7	8	AMY	IT	45.0	KANYAKUMARI
8	9	ANNIE	EEE	80.0	CHENNAI
9	10	ESTHER	ECE	90.0	TRICHY

In [93]:



```
#Delete the column "MARK"  
data.drop(['MARK'], axis=1)
```

Out[93]:

	Regno	NAME	DEPARTMENT	LOCATION
0	1	MAY	CSE	CHENNAI
1	2	ANDY	EEE	TRICHY
2	3	ANDREW	ECE	MADURAI
3	4	MICHAEL	ECE	THENI
4	5	MARY	EEE	COIMBATORE
5	6	JOY	CSE	VELLORE
6	7	ROY	CSE	THANJAVUR
7	8	AMY	IT	KANYAKUMARI
8	9	ANNIE	EEE	CHENNAI
9	10	ESTHER	ECE	TRICHY

In [94]:



```
#Display the Mark>80 and <95
print(data[(data['MARK']>80) & (data['MARK']<95)])
```

	Regno	NAME	DEPARTMENT	MARK	LOCATION
4	5	MARY	EEE	85.0	COIMBATORE
5	6	JOY	CSE	90.0	VELLORE
9	10	ESTHER	ECE	90.0	TRICHY

In [97]:



```
#Insert new values in 10th row.
data.loc[9] = {'Regno': 11, 'NAME': 'EMILY', 'DEPARTMENT': 'ECE', 'MARK': 87, 'LOCATION': 'CHE'}
data
```

Out[97]:

	Regno	NAME	DEPARTMENT	MARK	LOCATION
0	1	MAY	CSE	98.0	CHENNAI
1	2	ANDY	EEE	54.0	TRICHY
2	3	ANDREW	ECE	71.0	MADURAI
3	4	MICHAEL	ECE	62.0	THENI
4	5	MARY	EEE	85.0	COIMBATORE
5	6	JOY	CSE	90.0	VELLORE
6	7	ROY	CSE	100.0	THANJAVUR
7	8	AMY	IT	45.0	KANYAKUMARI
8	9	ANNIE	EEE	80.0	CHENNAI
9	11	EMILY	ECE	87.0	CHENNAI

In [103]:

#Insert new column as GPA.

GPA=[9.1,8.5,7.5,8.5,9.5,7.5,8.6,9.9,8.7,8.6]

data['GPA']=GPA

data

Out[103]:

	Regno	NAME	DEPARTMENT	MARK	LOCATION	GPA
0	1	MAY	CSE	98.0	CHENNAI	9.1
1	2	ANDY	EEE	54.0	TRICHY	8.5
2	3	ANDREW	ECE	71.0	MADURAI	7.5
3	4	MICHAEL	ECE	62.0	THENI	8.5
4	5	MARY	EEE	85.0	COIMBATORE	9.5
5	6	JOY	CSE	90.0	VELLORE	7.5
6	7	ROY	CSE	100.0	THANJAVUR	8.6
7	8	AMY	IT	45.0	KANYAKUMARI	9.9
8	9	ANNIE	EEE	80.0	CHENNAI	8.7
9	11	EMILY	ECE	87.0	CHENNAI	8.6

In [104]:

#Find the sum of null values

data.isnull().sum()

Out[104]:

```

Regno      0
NAME       0
DEPARTMENT 0
MARK       0
LOCATION     0
GPA        0
dtype: int64

```

In [108]:

#Create subset for first 3 columns

sub=data.loc[[4,6,9]]

sub

Out[108]:

	Regno	NAME	DEPARTMENT	MARK	LOCATION	GPA
4	5	MARY	EEE	85.0	COIMBATORE	9.5
6	7	ROY	CSE	100.0	THANJAVUR	8.6
9	11	EMILY	ECE	87.0	CHENNAI	8.6

In [109]:



```
data.head(14)
```

Out[109]:

	Regno	NAME	DEPARTMENT	MARK	LOCATION	GPA
0	1	MAY	CSE	98.0	CHENNAI	9.1
1	2	ANDY	EEE	54.0	TRICHY	8.5
2	3	ANDREW	ECE	71.0	MADURAI	7.5
3	4	MICHAEL	ECE	62.0	THENI	8.5
4	5	MARY	EEE	85.0	COIMBATORE	9.5
5	6	JOY	CSE	90.0	VELLORE	7.5
6	7	ROY	CSE	100.0	THANJAVUR	8.6
7	8	AMY	IT	45.0	KANYAKUMARI	9.9
8	9	ANNIE	EEE	80.0	CHENNAI	8.7
9	11	EMILY	ECE	87.0	CHENNAI	8.6

VISUALIZATION OF DATASET

#PART B QUESTION 14

Let us assume that the dataset consists of five columns namely "Rollno, Name, Dept, Mark and Location". Write the code for the following function

- How to upload the dataset?.
- Visualize the students mark using Line chart?.
- Display the Name and mark information using Bar chart
- Apply stacked bar chart for any one column
- Locate the legend in top left location
- Draw the scatter plot for any two columns
- Draw the line chart for any two columns
- Draw the Area chart for any two columns
- Display the x-axis and y-axis label.

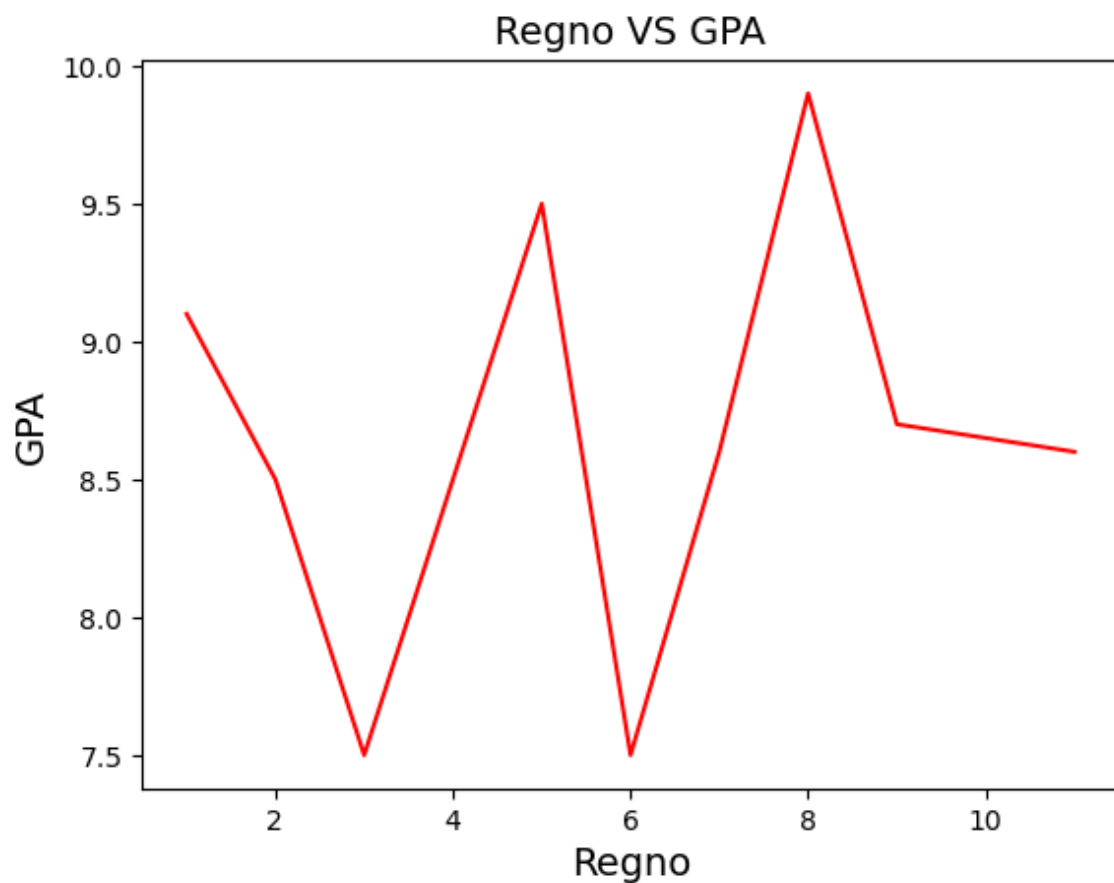
In [114]:



#PART B QUESTION 14

#Visualize the students mark using Line chart
#Draw the line chart for any two columns

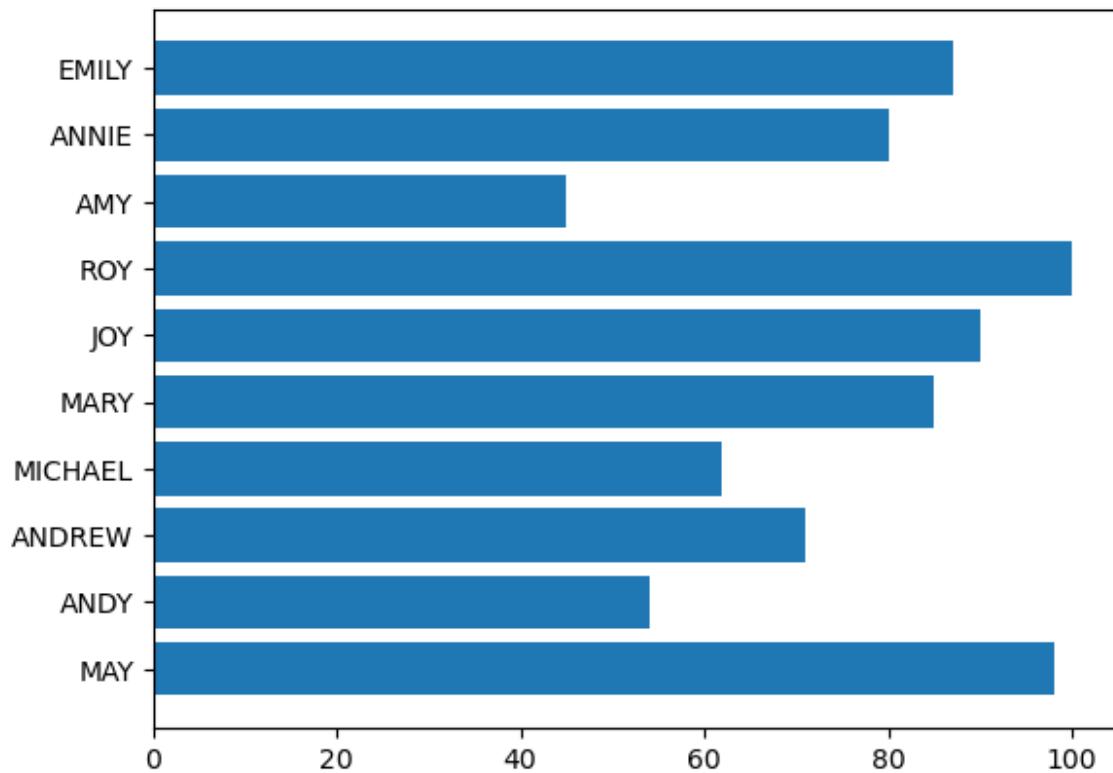
```
import matplotlib.pyplot as plt
plt.plot(data['Regno'], data['GPA'], color='red')
plt.title('Regno VS GPA', fontsize=14)
plt.xlabel('Regno', fontsize=14)
plt.ylabel('GPA', fontsize=14)
plt.show()
```



In [115]:

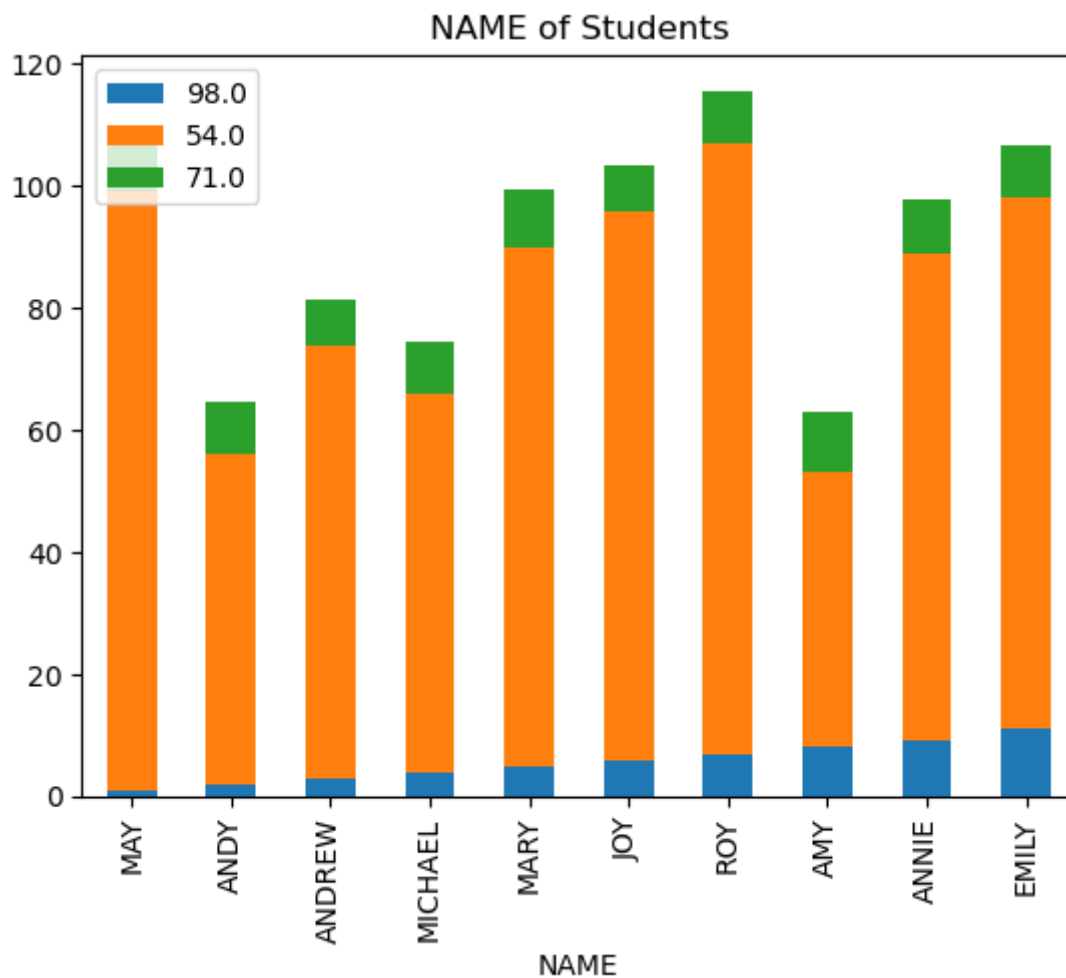


```
#Display the Name and mark information using Bar chart  
plt.barh(data['NAME'], data['MARK'])  
plt.show()
```



In [123]:

```
#Apply stacked bar chart for any one column  
data.plot.bar(x='NAME', stacked=True, title='NAME of Students')  
#Locate the legend in top left location  
plt.legend(data['MARK'], loc='upper left')  
plt.show()
```



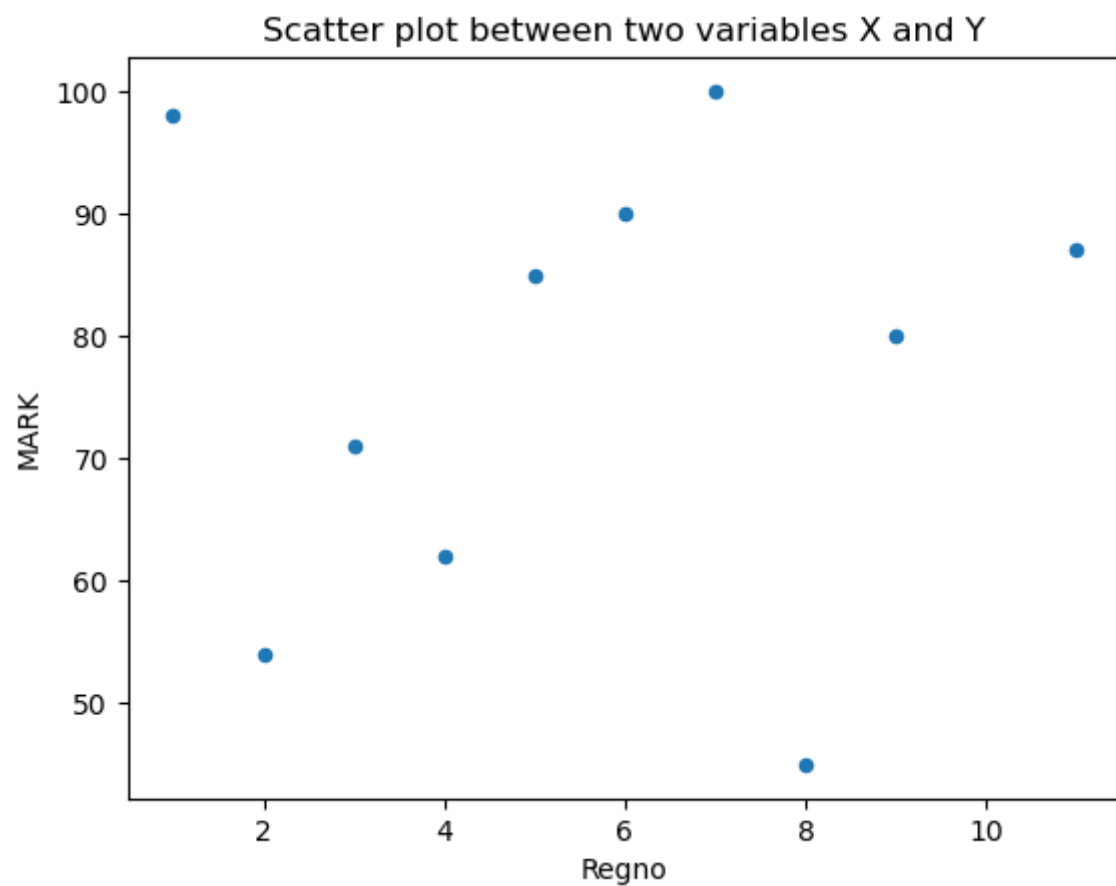
In [124]:



```
#Draw the scatter plot for any two columns
```

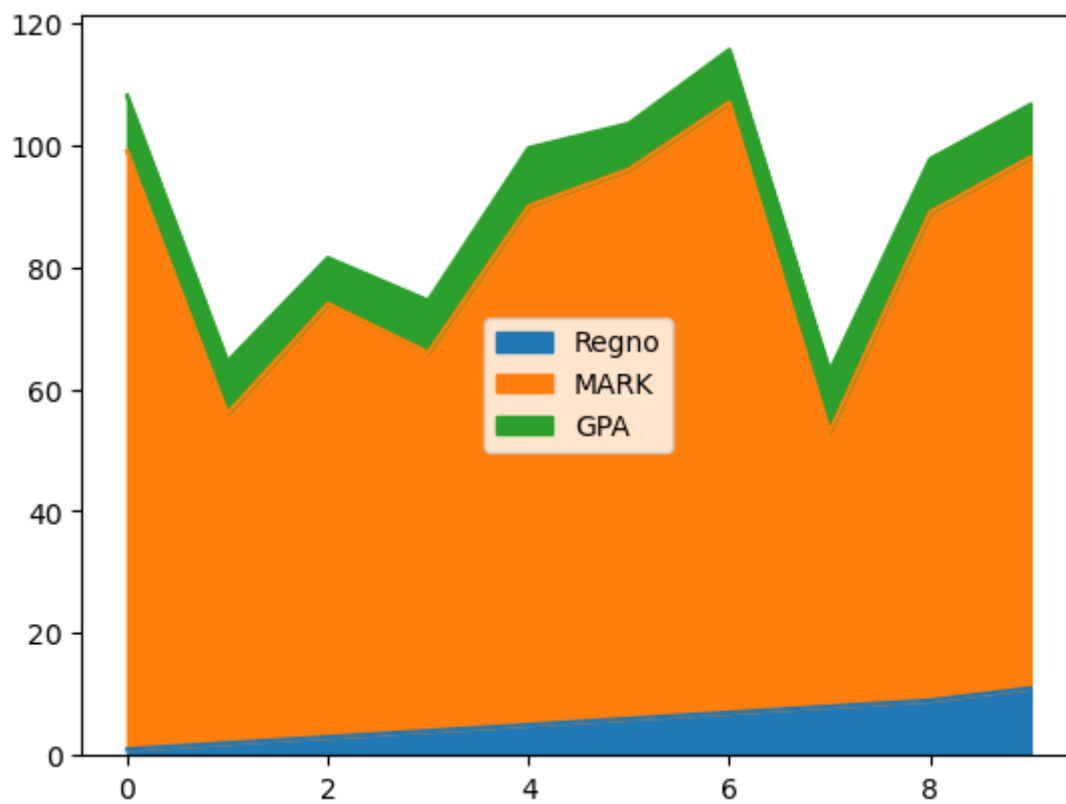
```
#Display the x-axis and y-axis label.
```

```
data.plot.scatter(x='Regno', y='MARK', title= "Scatter plot between two variables X and Y")
```



In [131]:

```
#Draw the Area chart for any two columns  
data.plot(kind='area', stacked=True)  
plt.show();
```



In []: